نشریه پژوهشهای علوم دامی ایران
جلد ۱۳، شماره ۱، بهار ۱۴۰۰، ص. ۱۶۰-۱۵۱

Iranian Journal of Animal Science Research
Vol. 13, No. 1, Spring 2021, p. 151-160

مقاله علمی- پژوهشی

شناسایی کل ریزماهواره‌ها در ژنوم شترهای دوکوهانه ایرانی با استفاده از داده‌های توالی‌یابی کل ژنوم

ناهیده زارع[۱]، نعمت هدایت ایوریق [۲*]، رضا سیدشریفی[۲]، رضا خلخالی[۳]، آزاده بوستان[۲]

تاریخ دریافت: ۱۳۹۸/۱۰/۲۴
تاریخ پذیرش: ۱۳۹۸/۱۲/۰۷

چکیده

شترهای دوکوهانه یکی از مقاومترین گونه‌های حیوانی در برابر شرایط سخت محیطی به شمار می‌روند که تعداد آنها در ایران در محدوده خطر قرار گرفته است. شناخت هر چه بهتر و دقیق‌تر این گونه، به خصوص در سطح مطالعات ژنومی، می‌تواند از طریق طراحی برنامه‌های مدیریت تنوع ژنتیکی، به حفظ این گونه کمک کند. هدف از انجام این مطالعه شناسایی ریزماهواره‌های ژنوم شترهای دوکوهانه ایرانی با استفاده از داده‌های توالی‌یابی کل ژنوم بود. در مطالعه حاضر از تعداد شش نفر شتر دوکوهانه متعلق به استان اردبیل خون‌گیری صورت گرفت. توالی‌یابی کامل ژنوم شترهای دوکوهانه با استفاده از پلتفرم ایلومینا و به صورت دو انتها (Paired-end) با اندازه ۱۰۰ جفت باز از هر طرف انجام شد. بعد از پالایش کیفی خوانش‌ها، گردآوری از نو آنها صورت گرفته و با استفاده از برنامه MISA به شناسایی تمام ریزماهواره‌های ژنوم‌های مورد مطالعه پرداخته شد. اندازه ژنوم‌های گردآوری شده برای شترهای دوکوهانه مورد مطالعه، در محدوده Gb ۱/۹۰ برای نمونه یک تا Gb ۱/۹۷ برای نمونه سه قرار داشت. مقدار N50 مربوط به کانتیگ‌های ایجاد شده برای شترهای دوکوهانه ایرانی از kb ۱۹/۱ برای نمونه یک تا kb ۵۱/۷ برای نمونه پنج متغیر بود. ریزماهواره‌های شناسایی شده در اندازه یک تا هشت نوکلئوتید بودند. کل ریزماهواره‌های شناسایی شده برای نمونه‌های شترهای دوکوهانه ایرانی در محدوده ۱۳۶۰۲۸ برای نمونه دو تا ۵۳۹۵۵۵ برای نمونه سه قرار داشت. همچنین با شناسایی ریزماهواره‌های هفت گونه پستاندار دیگر در این مطالعه، به مقایسه نتایج به دست آمده آنها با شتر دوکوهانه ایرانی پرداخته شد.

واژه‌های کلیدی: ایران، توالی‌یابی کل ژنوم، ریزماهواره، شتر دوکوهانه.

مقدمه

شترهای تک‌کوهانه و دوکوهانه در محیط‌های بیابانی آسیا و آفریقا زندگی می‌کنند و سازگاری آن‌ها با شرایط خشک شامل تحمل دمای بیش از ۴۰ درجه سلسیوس و از دست دادن آب بدن بیش از ۲۵ درصد کل وزن بدن می‌باشد (۱۹) در حالی‌که پستانداران غیر بیابانی با از دست دادن بیش از ۱۵ درصد وزن بدنشان دچار مرگ می‌شوند (۱۳). در مقابل، نزدیک‌ترین خویشاوندان شترها (Lamini) که فاقد کوهان بوده و شامل دو گونه وحشی (گواناکو و ویکوانا) و دو گونه اهلی (لاما و آلپاکا) هستند، در ارتفاعات آمریکای جنوبی زندگی می‌کنند و سازگاری مشابهی با محیط‌های بیابانی گرم ندارند. دمای بدن شترها ممکن است در طول روز بین ۳۴ تا ۴۱ درجه سلسیوس در نوسان باشد (۲۰). سطح قند خون در شترها دو برابر سایر نشخوارکنندگان است (۳). همچنین شترها مصرف زیاد نمک را تحمل می‌کنند، به طوری‌که مصرف نمک در آن‌ها هشت برابر بیشتر از گاو-ها و گوسفندها (۴) می‌باشد و با این حال دچار دیابت و فشار خون بالا نمی‌شوند. کشور ایران در یکی از خشک‌ترین مناطق کره زمین قرار گرفته است و لذا کمبود منابع آبی از یک سو و توانایی‌های ویژه شترها از سویی دیگر، می‌تواند آنها را به حیوانات ارزشمندی در زمینه تولید پروتئین در سطح کشور تبدیل کند. به نظر می‌رسد مطالعه این حیوانات به خصوص در سطح ژنوم، می‌تواند در حفظ این گونه و همچنین مدیریت تنوع ژنتیکی و تولیدی آنها موثر و مفید باشد.

ریزماهواره‌ها به عنوان بخشی از توالی‌های تکراری در ژنوم

۱- دانش آموخته کارشناسی ارشد ژنتیک و اصلاح دام و طیور-گروه علوم دامی، دانشکده کشاورزی و منابع طبیعی- دانشگاه محقق اردبیلی، اردبیل، ایران.
۲- دانشیار ژنتیک و اصلاح دام و طیور، گروه علوم دامی، دانشکده کشاورزی و منابع طبیعی- دانشگاه محقق اردبیلی، اردبیل، ایران.
۳- دکتری ژنتیک و اصلاح دام و طیور- دانشگاه محقق اردبیلی، اردبیل، ایران.
(*- نویسنده مسئول: (Email: nhedayat@uma.ac.ir
Doi:10.22067/ijasr.v13i1.85088

یوکاریوت‌ها، همچنین به عنوان تکرارهای توالی ساده یا SSR[1] ها نیز شناخته می‌شوند، تکرارهای پشت سر هم کوتاه که از موتیف‌های ۱ تا ۶ نوکلئوتیدی DNA تشکیل شده‌اند (۱۶). آنها بخش قابل توجهی از ژنوم موجودات پیچیده را تشکیل می‌دهند و اغلب نسبت بیشتری در مقایسه با توالی‌های کد کننده پروتئین را به خود اختصاص می‌دهند (۱۵). ریزماهواره‌ها در شکل‌دهی حدود ۳ درصد از ژنوم انسان شرکت داشته (۲۰) و توزیع غیر تصادفی را در بسیاری از ژنوم‌ها نشان می‌دهند (۲۳). توالی های ریزماهواره‌ای حاوی اطلاعات مفیدی هستند و به طور گسترده‌ایی برای ارزیابی تنوع ژنتیکی در داخل و بین جمعیت‌ها و همچنین برای بررسی تکامل در بین گونه-های دامی استفاده می‌شوند (۱۱). ریزماهواره‌ها (۱ تا ۱۰ نوکلئوتید) و ماهوارک‌ها (بیشتر از ۱۰ نوکلئوتید)، زیرشاخه‌های تکرارهای پشت سر هم (TR[2]ها) هستند که همراه با عناصرمتحرک[3]، نواحی تکراری ژنوم را تشکیل می‌دهند (۱۸). تکرارهای پشت سر هم به علت بی‌ثباتی‌شان با تکامل موجودات ارتباط دارند، به طوریکه با نرخ تغییر $۱۰^{۳}$ تا $۱۰^{۶}$ در هر نسل سلولی جز متغیرترین بخش‌های ژنومی به شمار می‌روند (۹). مطالعه این بخش از ژنوم می‌تواند در تعیین میزان تنوع ژنتیکی شترهای دوکوهانه و مقایسه آنها با دیگر گونه‌ها به لحاظ تکاملی، حائز اهمیت باشد. هدف ما از این مطالعه، شناسایی ریزماهواره‌های گسترده در ژنوم شترهای دوکوهانه ایرانی با استفاده از داده‌های توالی‌یابی کامل ژنوم بود و همچنین مقایسه آنها با پستانداران دیگر نیز از اهداف مطالعه حاضر بود.

مواد و روش

نمونه‌گیری و استخراج DNA

جهت شناسایی و بررسی نشانگرهای ریزماهواره‌ای در ژنوم شترهای دوکوهانه، از تعداد ۶ نفر شتر دوکوهانه‌ی ایرانی مربوط به استان اردبیل خونگیری انجام پذیرفت. خونگیری از رگ ورید گردنی و توسط ونوجکت ۴ میلی‌لیتری حاوی EDTA صورت گرفت. پس از خونگیری، نمونه‌ها بلافاصله به آزمایشگاه دانشکده کشاورزی دانشگاه محقق اردبیلی منتقل شده و تا زمان استفاده در دمای ۲۰- درجه سلسیوس نگهداری شدند. به منظور استخراج DNA نمونه‌ها، از کیت GeanAll (کره جنوبی) استفاده شد. پس از استخراج کل DNA از نمونه‌ها، کنترل کمّی DNA استخراجی با کمک دستگاه نانودراپ صورت پذیرفت. نسبت جذب طول موج‌های ۲۶۰ به ۲۸۰ برای نمونه DNAهای شترهای دوکوهانه مورد مطالعه، ۱/۹۶ بود. همچنین به منظور کنترل کیفی نمونه‌های DNA استخراجی، از ژل آگاروز یک درصد استفاده شد.

توالی‌یابی کل ژنوم

توالی‌یابی کامل ژنوم شترهای دوکوهانه با استفاده از پلتفرم ایلومینا و به صورت دوانتها[4] و با اندازه ۱۰۰ جفت‌باز از هر طرف انجام شد. از نرم‌افزار FastQC (www.bioinformatics.babraham.ac.uk/projectd/fastqc) برای کنترل کیفیت خوانش‌های توالی‌های خام استفاده شد. بررسی کیفیت حاکی از عدم وجود آلودگی‌های آداپتوری در خوانش‌های توالی‌های خام بود. با این حال، به منظور افزایش کیفیت داده‌ها، پالایشی بر روی این خوانش‌ها انجام گرفت. جهت پالایش خوانش-های خام از الگوریتم SLIDINGWINDOW (۴:۲۰) در نرم‌افزار Trimmomatic v0.36 (۶) استفاده شد. این الگوریتم با تقسیم خوانش‌ها به چندین قسمت و ارزیابی نمره کیفی این قسمت‌ها و حذف قسمت‌های پایین‌تر از این نمره کیفی اجرا می‌شود. در هنگام استفاده از این نرم‌افزار دستورهایی شامل LEADING، TRAILING و MINLEN به نرم‌افزار داده شد. در دستور LEADING بازهای دارای کیفیت پایین‌تر از آستانه تعریف شده برای برنامه از ابتدای خوانش‌ها حذف شدند و در دستور TRAILING بازهای دارای کیفیت پایین‌تر از آستانه تعریف شده برای برنامه از انتهای خوانش‌ها حذف شدند که حد آستانه تعریف شده برای هر دو این دستورات ۵ بود. دستور MINLEN تعریف شده در این مطالعه ۴۰ بود بدین معنی که اگر طول خوانش‌ها بعد از پالایش، به کمتر از ۴۰ جفت باز می‌رسیدند، توسط برنامه حذف می‌شدند. بعد از انجام این مراحل خوانش‌ها مجدداً توسط برنامه FastQC تحت کنترل کیفیت قرار گرفت.

گردآوری از نو[5] ژنوم شترهای دوکوهانه

گردآوری از نو خوانش‌های پالایش شده شترهای دوکوهانه با استفاده نسخه ۱۱ نرم افزار CLC Genomics Workbench (CLC Bio, Aarhus, Denmark) صورت پذیرفت. در واقع گردآوری از نو، سرهم‌بندی تعدادی توالی بدون استفاده از ژنوم مرجع می‌باشد. الگوریتمی که این نرم‌افزار به منظور گردآوری از نو استفاده می کند، گراف de Bruijn می‌باشد. لازم به ذکر است که گردآوری از نو در این نرم افزار در دو مرحله صورت می گیرد که مرحله اول ساخت کانتیگ‌ها با استفاده از خوانش‌های موجود و مرحله دوم همردیف کردن خوانش‌های ورودی با کانتیگ‌های ساخته شده است. پارامترهای استفاده شده در این مطالعه برای گردآوری خوانش‌ها

1 Single Sequence Repeats
2- Tandem Repeats
3- Transposable elements

4- Paired-end
5- de novo assembly

عبارتند از: ۳ برای هزینه عدم تطابق[1]، ۳ برای جریمه حذف و درج[2]، ۰/۵ برای length fraction و ۰/۸ برای similarity fraction.

شناسایی ریزماهواره های گسترده ژنوم

از نرم‌افزار MISA (MIcroSAtellite identification tool) (۵) برای شناسایی ریزماهواره‌های کل ژنوم های مورد مطالعه استفاده شد. این برنامه می‌تواند ریزماهواره‌های کامل و مرکب را شناسایی کند. ریزماهواره‌های مرکب در واقع بخشی از ریزماهواره‌ها هستند که توسط چندین نوکلئوتید در بین آنها فاصله ایجاد شده است. حداقل تکرارهای در نظر گرفته شده برای شناسایی ریزماهواره های تک نوکلئوتیدی حداقل دوازده تکرار، برای دو نوکلئوتیدی‌ها حداقل شش تکرار، برای سه نوکلئوتیدی‌ها و چهار نوکلئوتیدی‌ها حداقل پنج تکرار، برای پنج نوکلئوتیدی‌ها و شش نوکلئوتیدی‌ها حداقل چهار تکرار و برای هفت نوکلئوتیدی‌ها و هشت‌نوکلئوتیدی‌ها حداقل سه تکرار بود. حداکثر فاصله‌ایی که برای ریزماهواره‌های مرکب در نظر گرفته شد نیز ۱۰۰ جفت باز بود. مراحل ذکر شده برای ژنوم پستاندارن دیگر از جمله انسان (شماره دسترسی Genbank: GCF_000001405.37)، اسب (شماره دسترسی Genbank: GCF_000002305.2)، گاو (شماره دسترسی GCF_000003055.6:Genbank)، گوسفند (شماره دسترسی Genbank: GCF_000298735.2)، آلپاکا (۲۴)، شترهای تک کوهانه (۱۶) و همچنین شتر دوکوهانه غیر ایرانی (۱۴) نیز صورت پذیرفته و نتایج آنالیزها در مورد این پستانداران با شترهای دوکوهانه ایرانی مورد مقایسه قرار گرفت.

نتایج و بحث

شکل ۱ به وضوح تاثیر پالایش کیفی بر روی افزایش کیفی خوانش‌های خام را نشان می‌دهد. محدوده‌های رنگی نشان دهنده کیفیت خوانش‌ها هستند. از پایین به بالا با تغییر رنگ محدوده‌ها کیفیت داده‌ها افزایش می‌یابند. اعداد محور افقی، تعداد بازها و محور عمودی نیز نمره کیفی آنها را نشان می‌دهد و خط آبی نیز نمودار کیفی است.

اندازه ژنوم گردآوری شده برای نمونه شترهای دوکوهانه مورد مطالعه در محدوده ۱/۹۰ Gb برای نمونه یک تا ۱/۹۷ Gb برای نمونه سه قرار داشت. اندازه ژنوم گردآوری شده برای شترهای دوکوهانه در مطالعه حاضر کمتر از مقادیر گزارش شده در مطالعه جیریموتو و همکاران (۱۴) (۲/۳۸ Gb)، وو و همکاران (۲۴) (Gb ۲/۰۱) و فیتاک و همکاران (۸) (۲/۰۶ Gb) بود. اندازه ژنوم نمونه پنج و شش با شترهای تک‌کوهانه ایرانی (۱/۹۴ Gb) (۱۶) برابر بودند. مقدار N50 مربوط به کانتیگ‌های ایجاد شده برای شترهای دو-کوهانه ایرانی از ۱۹/۱ kb برای نمونه یک تا ۵۱/۷ kb برای نمونه پنج متغیر بود (جدول۱). معیار N50 یکی از پارامترهای اندازه‌گیری کیفی گردآوری ژنوم می‌باشد و اندازه بزرگتر به معنی گردآوری بهتر خوانش‌ها به شمار می‌رود. در واقع عدد N50 بیانگر این مسئله می-باشد که به عنوان مثال کانتیگ‌های ۴۹ هزار بازی یا بزرگتر بیش از ۵۰ درصد ژنوم گردآوری شده در نمونه شش را تشکیل می دهند. اندازه N50 در این مطالعه در نمونه‌های پنج و شش بزرگتر از شتر تک‌کوهانه آفریقایی (۴۰/۲ kb؛ فیتاک و همکاران (۸)) و در نمونه-های دو و چهار بزرگتر از شتر نژاد تارگویی (Genbank ؛ ۳/۵ kb) بود؛ در حالیکه از مقادیر گزارش شده برای تک‌کوهانه عربی(kb ۶۹/۱؛ وو و همکاران(۲۴))، دوکوهانه وحشی (۹۰/۳ kb؛ جیریموتو و همکاران(۱۴)) و شترهای تک‌کوهانه ایرانی (۵۴/۶ kb و ۵۴/۰۳؛ خلخالی و همکاران (۱۶)) کوچک‌تر بود.

توالی‌های ریزماهواره‌ای یکی از متنوع‌ترین قسمت‌های ژنوم موجودات مختلف به شمار می روند (۷). این توالی‌ها به صورت تصادفی در سرتاسر ژنوم بیشتر یوکاریوت‌ها پراکنده شده و به مقدار بالایی دارای چند شکلی هستند (۱). به همین دلایل، توالی های ریزماهواره‌ای به طور گسترده‌ای برای مطالعه تنوع ژنتیکی در داخل و بین جمعیت‌های گونه‌های مختلف دامی از جمله گاو (۲۲)، گوسفند(۱۲)، بز (۱۰) و اسب (۲) مورد استفاده قرار می‌گیرند.

در این مطالعه توالی‌های ریزماهواره‌ای شترهای دوکوهانه ایرانی و نیز شترهای تک‌کوهانه ایرانی، تک کوهانه غیر ایرانی، شتر دوکوهانه، آلپاکا، گاو، گوسفند، اسب و انسان شناسایی و با یکدیگر مقایسه شدند. ریزماهواره‌های شناسایی شده در اندازه یک تا هشت نوکلئوتید بودند. کل ریزماهواره‌های شناسایی شده برای نمونه‌های شترهای دوکوهانه ایرانی در محدوده ۱۳۶۰۲۸ جفت‌بازی برای نمونه دو تا ۵۳۹۵۵۵ جفت‌بازی برای نمونه سه قرار داشت. نتایج به دست آمده نشان می‌دهد که در نمونه‌های یک، دو، سه، چهار، پنج و شش به ترتیب حدود ۳/۱۳ Mb، ۲/۳۵ Mb، ۹/۲۶ Mb، ۷/۱ Mb، Mb ۸/۹۹ و ۸/۸۶ Mb ژنوم شترهای دوکوهانه از توالی‌های ریزماهواره‌ای تشکیل شده است (جدول ۲). لازم به ذکر است که شناسایی مقادیر متفاوت ریزماهواره‌های شناسایی شده در ژنوم شترهای مورد مطالعه، مربوط به میزان کیفیت گردآوری آنها می‌باشد. به طور مشخص، ژنوم‌هایی که با کیفیت بالاتری گردآوری شده‌اند، منجر به شناسایی ریزماهواره‌های بیشتری شده‌اند. در پستاندارنی که در این مطالعه بررسی شد، انسان با ۲۵/۷ Mb و اسب با ۷/۸۱ Mb دارای بیشترین و کمترین اندازه کل ریزماهواره‌ها بودند.

1-mismatch cost
2-insertion and deletion cost

جدول ۱–آماره‌های مربوط به گردآوری از نو ژنوم شترهای دوکوهانه ایرانی

Table 1– Summary of the Iranian Bactrian camels' genome assembly

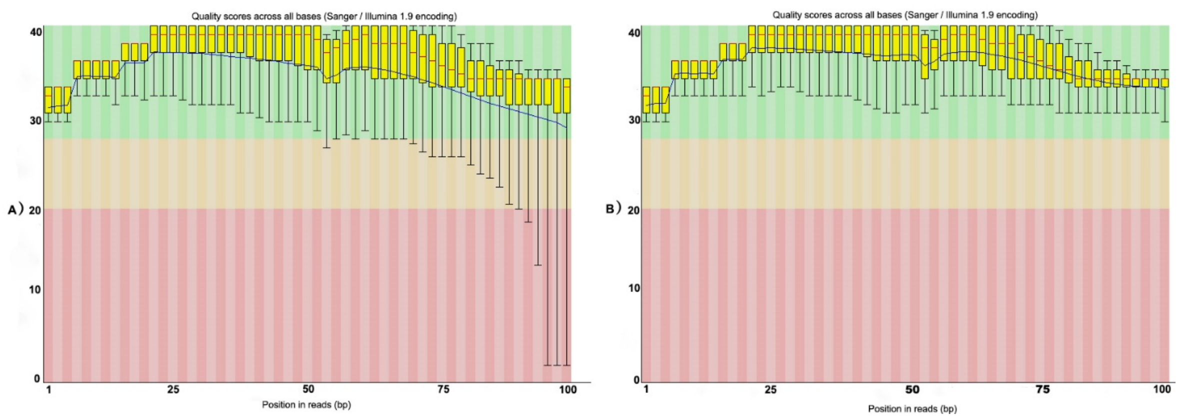
کانتیگ‌ها Contigs	نمونه یک Sample one	نمونه دو Sample two	نمونه سه Sample three	نمونه چهار Sample four	نمونه پنج Sample five	نمونه شش Sample six
طول N50 N50 length (bp)	19123	39174	21927	34156	51746	49043
میانگین طول کانتیگ Average contig length (bp)	103359	18512	6885	1408	12972	12109
اندازه ژنوم Genome size (Gb)	1.90	1.92	1.97	1.96	1.94	1.94

جدول ۲– نتایج شناسایی توالی‌های ریزماهواره‌ای در شترهای دوکوهانه ایرانی و هفت گونه پستاندار دیگر

Table 2– Results of the identification of SSRs in Iranian Bactrian camels and seven other mammal species

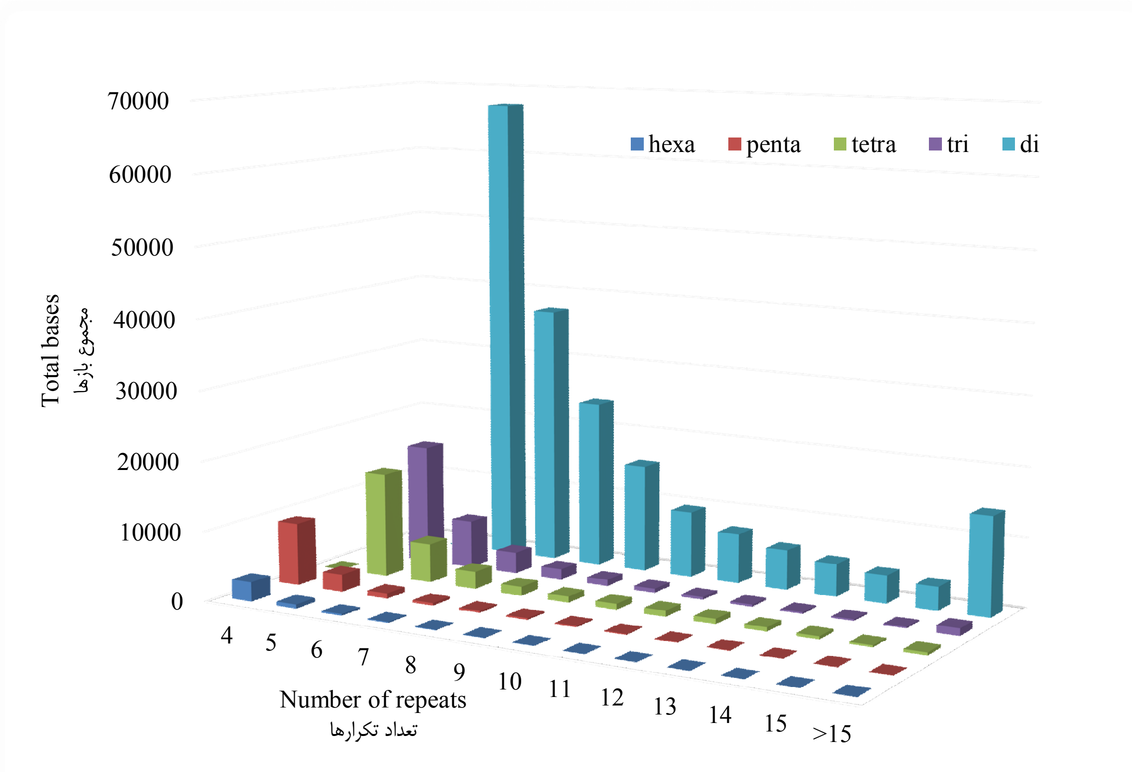
گونه Species	یک Mono	دو Di	سه Tri	چهار Tetra	پنج Penta	شش Hexa	هفت Hepta	هشت Okta	مجموع Total (bp)
تک‌کوهانه یزدی Yazdi dromedary (YaD)	3670743	3429832	638283	781068	329445	111366	97839	59368	9117944
تک‌کوهانه طرودی Trodi dromedary (TrD)	3669900	3417918	636564	778964	325930	110706	97790	58208	9095980
تک‌کوهانه عربی Arabian dromedary	4121293	3859778	685755	1390196	401970	129516	97986	89456	10775950
شتردوکوهانه Bactrian camel	4082129	3791158	615105	1147212	366725	125718	99785	92632	10320464
آلپاکا Alpaca	2065967	4102030	627525	1100084	348420	123546	152278	72032	8591882
گاو Cattle	5045498	4858676	2046501	254296	1537845	54444	94374	26576	13918210
گوسفند Sheep	4657293	4889182	1770183	392980	1589625	62922	78764	57312	13498261
اسب Horse	3344142	3025388	577437	532720	180225	47934	64001	37848	7809695
انسان Human	12287665	6754988	1537530	3001280	1439420	291282	304525	149072	25765762
نمونه یک Sample one	1290139	1145018	215199	275028	108770	40746	34503	21112	3130515
نمونه دو Sample two	951921	866386	169170	209280	85655	30432	29225	17288	2359357
نمونه سه Sample three	3822194	3384294	633267	827776	328280	112908	97286	61008	9267013
نمونه چهار Sample four	2925948	2586926	495441	624620	252680	89394	78624	46984	7100617
نمونه پنج Sample five	3602706	3381812	637404	775116	328470	111294	99225	59192	8995219
نمونه شش Sample six	3567986	3327558	630276	752264	323115	107376	97370	57968	8863913

با توجه به شکل ۲، نتایج نشان می‌دهد که پراکنش ریزماهواره‌ها با افزایش اندازه تعداد توالی‌ها کاهش می‌یابد به طوری که توالی‌های یک و دو بازی بیشترین فراوانی را دارند. در حالی که توالی‌های پنج و شش بازی کمترین پراکنش ریزماهواره‌ها را به خود اختصاص دادند. همچنین تعداد ریزماهواره‌های با واحدهای تکرار شونده بالا نسبت به ریزماهواره‌های با واحدهای تکرار شونده پایین، کمتر می‌باشند.



شکل ۱– کیفیت خوانش‌های مربوط به یکی از شترهای دوکوهانه ایرانی قبل (A) و بعد (B) از پالایش کیفی

Figure 1–Quality of reads related to one of the Iranian Bactrian camels before (A) and after (B) of quality trimming



شکل ۲– پراکنش ریزماهواره‌های با اندازه مختلف در سطح ژنوم شترهای دوکوهانه ایرانی

Figure 2- Distribution of different SSR motifs in Iranian Bactrian camel's genome

جدول ۳، ده ریزماهواره با بیشترین تکرار در هفت گونه را نشان می‌دهد. بیش از ۷۴ درصد ریزماهواره‌های شناسایی شده به ده ریزماهواره با تکرار بالا در همه هفت گونه مربوط می‌باشد. موتیف T در نمونه شترهای دوکوهانه یک و شش، شترهای تک‌کوهانه ایرانی، دوکوهانه، گاو، گوسفند، اسب و انسان فراوان‌ترین موتیف است. در نمونه‌های دو، سه، چهار، پنج، تک‌کوهانه غیر ایرانی و آلپاکا فراوان‌ترین موتیف، A می‌باشد. در تمام نمونه‌ها به جز گاو، گوسفند و اسب AC فراوان‌ترین موتیف دو بازی است و TG در سه گونه ذکر

شده و در میان دو بازی‌ها، رتبه اول را دارا می‌باشد. در نمونه شماره دو موتیف C جزو ده ریزماهواره با بالاترین تکرار است که در هیچ کدام از گونه‌های دیگر مشاهده نمی‌شود. در ژنوم گاو و گوسفند به ترتیب دو ریزماهواره سه بازی AGC ، TGC و یک ریزماهواره سه بازی AGC و یک ریزماهواره پنج بازی ACTGA در فهرست ده ریزماهواره مشاهده می‌شود. بیشترین فراوانی در بین کلاس‌های مختلف ریزماهواره‌ها در تمام شش نمونه شتر دوکوهانه در تک بازی-ها A/T، در دو بازی‌ها AC/GT، در سه بازی‌ها AAT/ATT، در چهار بازی‌ها AAAC/GTTT، در پنج بازی‌ها AAAAC/GTTTT، در شش بازی‌ها AAAAAC/GTTTTT، در هفت بازی‌ها AAAAAAC/GTTTTTو در هشت بازی‌ها AAAAAAAC/GTTTTTTTبود. بررسی الگوهای مختلف ریزماهواره‌ها در گونه‌های مختلف می‌تواند در بحث تکامل و مدیریت تنوع ژنتیکی این گونه‌ها کمک کننده باشد.

جدول۳-ده توالی ریزماهواره‌ای با بالاترین تکرار در شترهای دوکوهانه ایرانی و هفت گونه دیگر

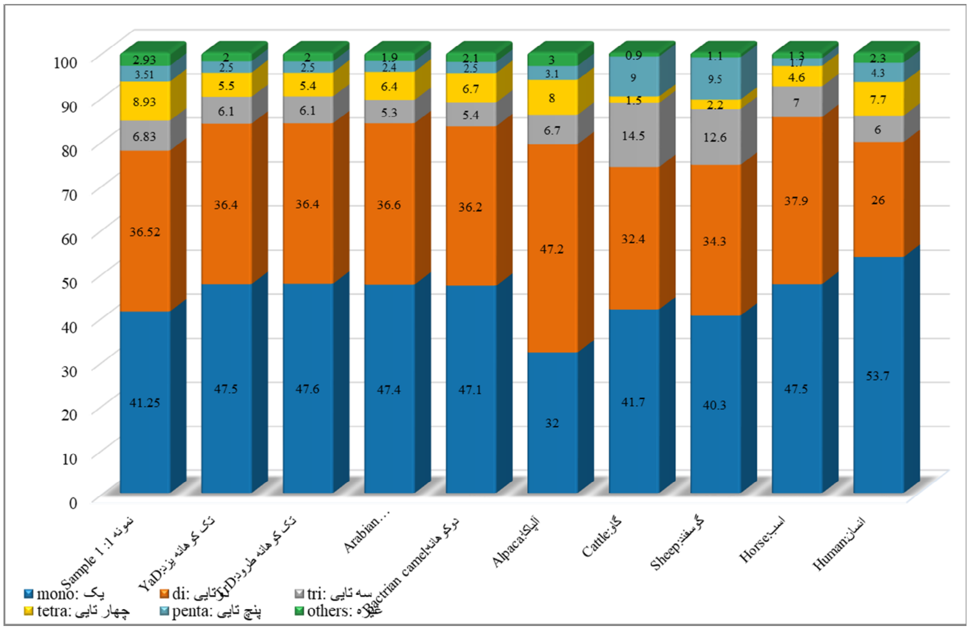
Table 3-Ten microsatellites with most frequency in Iranian Bactrian camel and other seven mammalian genome

ردیف Rank	یزدی YaD	دوکوهانه Bactria	آلپاکا Alpaca	گاو Cattle	گوسفند Sheep	اسب Horse	انسان Human	نمونه یک Sampleone	نمونه دو Sampletwo	نمونه سه Sample three	نمونه چهار Sample four	نمونه پنج Sample five	نمونه شش Sampel six
1	T	T	A	T	T	T	T	T	A	A	A	A	T
2	A	A	T	A	A	A	A	A	T	T	T	T	A
3	AC	AC	AC	TG	TG	TG	AC	AC	AC	AC	AC	AC	AC
4	TG	TG	TG	AC	AC	AC	TG	TG	TG	TG	TG	TG	TG
5	AT	AT	AT	AT	AT	TA	AT	GT	GT	AT	AT	AT	AT
6	TA	TA	TA	AGC	TA	CA	TA	AT	CA	TA	GT	TA	TA
7	GT	GT	GT	TA	CA	AT	GT	CA	AT	GT	CA	GT	GT
8	CA	CA	CA	CA	GT	GT	CA	TA	TA	CA	TA	CA	CA
9	TC	TC	TC	GT	AGC	TC	TC	TC	TC	TC	TC	TC	TC
10	AG	AG	AG	TGC	ACTGA	AG	AG	AG	C	AG	AG	AG	AG
From all (%)	78.8	79	74.9	76.1	74.2	78.5	77.3	77.8	76.6	78.5	77.9	79.1	79.2

شکل ۳، درصد ریزماهوراه‌های با اندازه مختلف در یکی از ژنوم شترهای دوکوهانه به عنوان نمونه و هفت گونه دیگر را نشان می‌دهد. درصد ریزماهواره‌ها نشان می‌دهد که تقریباً در تمام گونه‌ها الگوی پراکنش ریزماهواره‌ها در توالی‌های یک و دو بازی تقریباً مشابه است. نکته جالب در مورد آلپاکا این است که علیرغم هم خانواده بودن با شترهای تک و دوکوهانه، الگوی متفاوتی از نظر تعداد ریزماهواره‌ها با اندازه مختلف را نشان داد. در گوسفند و گاو درصد ریزماهواره‌های سه بازی بیشتر از گونه‌های دیگر است. اما درصد ریزماهواره‌های چهار بازی کمتر از گونه‌های دیگر می‌باشد. به طور کلی گوسفندها و گاوها الگوی مشابهی را نشان دادند که این مورد درباره شترهای تک و دوکوهانه با هم نیز صدق می‌کرد. این تصویر در واقع می تواند تابلویی ابتدایی از مسیر متفاوت طی شده تکامل در گونه های مورد مطالعه باشد که مطالعات عملکردی ریزماهواره‌ها با اندازه های مختلف می تواند به کامل شدن این تابلو کمک کرده و اطلاعات دقیق‌تری از مسیر تکاملی در این پستانداران را در اختیار قرار دهد.

نتیجه‌گیری کلی

نتایج به دست آمده در مطالعه کنونی که مجموعه‌ای از نشانگرهای ریزماهواره‌ای شناسایی شده در سطح ژنوم شترهای دوکوهانه ایرانی است را می‌توان برای طراحی مطالعات تنوع ژنتیکی به کار بست. شناخت سطح تنوع ژنتیکی با استفاده از مطالعات ژنومیک می‌تواند در شناخت وضعیت یک گونه و در نهایت مدیریت حفظ و نگهداری آن گونه حیوانی کمک کننده باشد. شترهای دوکوهانه ایرانی در وضعیت نامناسبی از نظر تعداد و پراکنش قرار داشته و عدم توجه به آنها می‌تواند این گونه ارزشمند حیوانی را بیش از پیش در معرض انقراض قرار دهد. لذا تجمیع اطلاعات حاصل از مطالعات به خصوص در سطح ژنوم می‌تواند منجر به طراحی برنامه‌هایی مناسب و موثر برای حفظ و نجات این گونه حیوانی در کشور شود.



شکل۳– درصد ریزماهواره‌های با اندازه مختلف در ژنوم شترهای دوکوهانه ایرانی و هفت پستاندار دیگر

Table 3- Distribution pattern of microsatellites in Iranian Bactrian camel and other seven mammalian genome

منابع

1- Abdul-Muneer, P. 2014. Application of microsatellite markers in conservation genetics and fisheries management: recent advances in population structure analysis and conservation strategies. Genetics Research International, 2014: 691759.

2- Ala-Amjad, M., H. Yeganeh, and M. Sadeghi. 2017. Study of Genetic variation in Iranian Kurdish horse using microsatellite marker. Iranian Journal of Animal Science, 48(3):342-335. (In Persian)

3- Al-Ali, A., H. Husayni, and D. A. Powe. 1988. Comprehensive biochemical analysis of the blood of the camel (Camelus dromedarius). Comparative Biochemistry and Physiology, 89(1): 35–37.

4- Ali, T. A. 1994. A Manual for the Primary Animal Health Care Worker. Food and Agriculture Organization of the United Nations (FAO).

5- Beier, S., T. Thiel, T. Münch, U. Scholz, and M. Mascher. 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics, 33(16):2583-2585.

6- Bolger A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15):2114-2120.

7- Ellegren, H. 2004. Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics, 5(6):435–445.

8- Fitak, R. R., E. Mohandesan, J. Corander, and P. A. Burger. 2016. The de novo genome assembly and annotation of a female domestic dromedary of North African origin. Molecular Ecology Resources, 16(1):314–324.

9- Gemayel, R., J. Cho, S. Boeynaems, and K. J. Verstrepen. 2012. Beyond Junk-Variable Tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. Genes, 3(3):461-480.

10- Guang-Xin, E., Q. H. Hong, Y. J. Zhao, Y. H. Ma, M. X. Chu, L. Zhu, and Y. F. Huang. 2019. Genetic diversity estimation of Yunnan indigenous goat breeds using microsatellite markers. Ecology and Evolution, 9(10): 5916.

11- Hampton, J. O., P. B. S. Spencer, D. L. Alpers, L. E. Twigg, A. P. Woolnough, J. Doust, T. Higgs, and J. Pluske. 2004. Molecular techniques, wildlife management and the importance of genetic population structure and dispersal: a case study with feral pigs. Journal of Applied Ecology, 41(4):735–743.

12- Huson, K. M., W. Haresign, M. Hegarty, T. Blackmore, C. Potter, and N. McEwan. 2015. Assessment of genetic relationship between six populations of Welsh Mountain sheep using microsatellite markers. Czech Journal of Animal Science, 60(5): 216-223.

13- Ingram, D. L., and L. E. Mount. 1975. Man and Animals in Hot Environments. Springer Science & Business Media.

14- Jirimutu, Z. W., G. Ding, G. Chen, Y. Sun, Z. Sun, H. Zhang, L. Wang, S. Hasi, Y. Zhang, J. Li, and Y. Shi. 2012.

Genome sequences of wild and domestic Bactrian camels. Nature Communications, 3: 1202.
15- Katti, M. V., P. K. Ranjekar, and V. S. Gupta. 2001. Differential distribution of simple sequence repeats in eukaryotic genome sequences. Molecular Biology and Evolution, 18(7):1161–1167.
16- Khalkhali-Evrigh, R.., S. H. Hafezian, N. Hedayat-Evrigh, A. Farhadi, and M. R. Bakhtiarizadeh. 2019. Genome-Wide Identification of Microsatellites and Transposable Elements in the Dromedary Camel Genome Using Whole-Genome Sequencing Data. Frontiers in Genetics, 10(2019):692.
17- Kelkar, Y. D., N. Strubczewski, S. E. Hile, F. Chiaromonte, K. A. Eckert, and K. D. Makova. 2010. What is a microsatellite: a computational and experimental definition based upon repeat mutational behavior at A / T and GT / AC repeats. Genome Biology and Evolution, 2: 620–635.
18- Leclercq, S., E. Rivals, and P. Jarne. 2010. DNA slippage occurs at microsatellite loci without minimal threshold length in humans: a comparative genomic approach. Genome Biology Evolution, 2:325–335.
19- Schmidt-Nielsen, K. 1959. The physiology of the camel. Scientific American, 201: 140–151.
20- Schmidt-Nielsen, K. 1964. Desert Animals. Physiological Problems of Heat and Water. Oxford University Press.
21- Subramanian, S., R. K. Mishra, and L. Singh. 2003. Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. Genome Biology, 4(2): 1-10.
22- Sun, W., C. Lei, X. Lei, and Y. Zhang. 2008. Genetic variation in eight Chinese cattle breeds based on the analysis of microsatellite markers. Genetics Selection Evolution, 40(6): 1-12.
23- Toth, G., Z. Gáspári, and J. Jurka. 2000. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Research, 10(7):967–981.
24- Wu, H., X. Guang, M. B. Al-Fageeh, J. Cao, S. Pan, H. Zhou, L. Zhang, M. H. Abutarboush, Y. Xing, and A. S. Alshanqeeti. 2014. Camelid genomes reveal evolution and adaptation to desert environments. Nature Communications, 5: 5188.

Iranian Journal of Animal Science Research
Vol. 13, No. 1, Spring 2021, p. 151-160

نشریه پژوهشهای علوم دامی ایران
جلد ۱۳، شماره ۱، بهار ۱۴۰۰، ص. ۱۶۰-۱۵۱

Identification of total microsatellites in the genome of Iranian Bactrian camels using whole genome sequencing data

N. Zare[1], N. Hedayat Evrigh[2*], R seyedsharifi[2], R. Khalkhali Evrigh[3], A. Boustan[2]
Submitted: 14-10-2020
Accepted: 26-02-2020

Introduction Bactrian camels are known as one of the resistant species to harsh environmental conditions. The camel's body temperature may vary from 34 to 41 °C throughout the day. They can survive if they lose body water greater than 25% of total body weight, while, in non-desert mammals, losses of greater than 15% are deadly. Since, Iran is located in one of the most arid regions of the world and water resources shortage, also special capabilities of camels, this species can be a valuable source of protein in the country. The study of genetic diversity is one of the most widely studies in domestic animals and microsatellites are widely used in this field. Microsatellite sequences contain useful information and are widely used to assess genetic diversity within and between populations, as well as to investigate the evolution process between species. The main aim of the present study was to identify the total microsatellites in the genome of Iranian Bactrian camels using whole genome sequencing data and compare them with other mammalians.

Materials and Methods This study was carried out to identify genome wide microsatellites on six Bactrian camels from Ardabil province. Blood samples were collected from the jugular vein using 4 ml vacutainer tubes and stored at -20C° until use. Illumina HiSeq 2000 technology (Illumina, USA) was used for whole genome sequencing of samples. Sequencing was performed using the paired-end method with 100 bp at both ends of the reads. The quality control of raw sequence reads was performed using FastQC software. The SLIDINGWINDOW (4:20) algorithm of Trimmomatic v0.36 program was used to quality filter of raw reads. After filtration of reads with low quality, reads shorter than 40 bp were discarded. The *de novo* assembly of trimmed reads from Bactrian camels was done using CLC Genomics Workbench 11 software (CLC Bio, Aarhus, Denmark). The parameters used in this study for *de novo* assembly of trimmed reads were: 3 for mismatch cost, 3 for deletion and insertion cost, 0.5 for length fraction, and 0.8 for similarity fraction. Assembled genomes were searched for identifying the microsatellites using MISA with motif size ranging from mono-nucleotide to octo-nucleotide. The minimum repeat numbers were defined as 12 for mono-, 6 for di-, 5 for tri- and tetra, 4 for penta- and hexa-, and 3 for hepta- and octo-nucleotide repeat SSRs. Microsatellite motifs that interrupted by 100 nucleotides were considered as compound microsatellites. Also, several mammalians assembled genomes were downloaded and searched for microsatellite loci, including Arabian dromedary camel, Bactrian camel, alpaca, horse, cattle, sheep, and human.

Results and Discussion The assembled genome size for the Bactrian camels were ranged from 1.90 for sample one to 1.97 for sample three. Also, the N50 length for the assembled contigs of Iranian Bactrian camels were ranged from 19.1 kb for sample one to 51.5 kb for sample five. The contig N50 length is one of the qualitative measurement parameters of genome assembly and a larger size means better assembly. The total microsatellites loci identified for Iranian Bactrian camels ranged from 136028 for sample two to 539555 for sample three. The results show that the

1-Graduted in Genetics and Animal Breeding, Department of Animal Science, Faculty of Agriculture and Natural Resource, University of Mohaghegh Ardabili.
2- Associate Professor of Genetics and Animal Breeding, Department of Animal Science, Faculty of Agriculture and Natural Resource, University of Mohaghegh Ardabili.
3- Genetics and animal breeding, PhD. University of Mohaghegh Ardabili.
(* Corresponding author: Email: nhedayat@uma.ac.ir)
Doi:10.22067/ijasr.v13i1.85088

genome of samples one, two, three, four, five and six contained 3.13 Mb, 2.35 Mb, 9.26 Mb, 7.1 Mb, 8.99 Mb and 8.86 Mb microsatellites, respectively. It should be noted that the difference in the microsatellites of SSRs in the Iranian Bactrian camel genomes is due to their different qualities in assembly. In mammals examined in this study, humans with 25.7 Mb and horses with 7.81 Mb had the highest and lowest total size of microsatellites, respectively. The results revealed that the number of microsatellites decreases with increasing in them, repeats, so that, one and two repeats sequences are the most frequent motifs. More than 74% of the identified microsatellites belong to the ten microsatellites with the highest number in all seven species. The motif T is the most frequent motif in the samples one and six Iranian Bactrian camels, Iranian dromedary camels, Bactrian camel, cattle, sheep, horses and humans. In samples two, three, four, five, the non-Iranian dromedary camel and alpaca motif A is the most abundant motif. The finding of this study will be applied as a valuable resource for further studies on camel breeding, especially on Iranian Bactrian camels. A large number of camel's SSR markers developed in this study established a valuable resource for the investigation of genetic diversity and may improve the development of breeding programs in Iranian Bactrian camels in the future.

Keywords: Bactrian camel, Iran, Microsatellite, Whole genome sequencing.